MINOS Computing/Offline Software Status Report

Liz Buckley-Geer
March 10, 2006
Report #5

Project Description

MINOS Computing
http://wwwserver2.fnal.gov/cfdocs/projectsdb/projdetail.cfm?ProjectID=81

MINOS Offline Software
http://wwwserver2.fnal.gov/cfdocs/projectsdb/projdetail.cfm?ProjectID=292

Summary

This status report covers the period from January 2006 to March 2006.

MINOS Computing

This task covers items such as data handling, control room support etc. The people involved are Liz Buckley-Geer, Art Kreymer and John Urish.

1. **MINOS Control Room**
   John Urish continues to support the computers in the control room on WH12NW. We kept up with the usual kernel updates. There were no off hours support issues in the last 3 months.
   During the accelerator shutdown we plan to upgrade all machines to SL4 and do some reconfiguration. John had made a plan for this work and we are waiting to get it scheduled.

2. **Hardware deployment**
   We have still not yet deployed the development Oracle server but we are getting closer. The deployment has been held up due to a shortage of personnel in the database support group.

3. **Data Handling**
   Art Kreymer is the primary data handling support person. SAM has been in production since September 2005. We ran 882 projects between January 1 and March 6.

   Enstore and Dcache underwent major upgrades during this period which caused some instabilities in MINOS operations. We now have dedicated write pools for the raw data. They are rather large compared to the size of our raw data so after the dust has settled on the upgrades we will do some tuning with the support people. There was a problem with a corrupted Dcache script which caused 5 files that the archiver had transferred to Dcache and had been marked as archived to not actually make it to tape. Luckily we were able to recover them from disk and re-archive them. There was one raw data file

that had already been deleted from the DAQ buffer but we located a copy on a control room machine.

We submitted a paper to CHEP on “Lightweight Deployment of SAM”.

We removed write access to PNFS from most of the nodes on the MINOS cluster after a couple of raw data files mysteriously disappeared.

The problem encountered in October 2004 with treatment of daylight savings time in SAM surfaced again. The data in the database is okay because we used the UTC time type in SAM to declare the data – this was the problem that we fixed back then. However the **sam get metadata** command messes the times up by one hour for files that were taken during the DST period. We also discovered that the times on the reco files are off by six hours. This is probably because we clone the reco entry from the raw file by getting the metadata from SAM in a python script and then declaring it with the necessary changes. This declare did not use the UTC time type so ran into the python features related to time zones etc. We need to fix **sam get metadata** so that it returns the correct time that is in the database. We store all times in UTC. We also need to fix the reco files.

We still have to incorporate the Monte Carlo data into SAM. We are a bit behind on this. Initially we were just going to declare the existing files with minimal metadata but we decided to define some appropriate parameters as well. This is in development at the moment. We also need to add the data from the CERN testbeam.

Tests have been made to concatenate the production output into more manageable units. This should lead to a reduction of 24 in the numbers of files produced by the reconstruction stage. Hopefully this can be put into place for the next major re-processing.

The ISA group will proceed with the reorganization of the MINOS PNFS database on March 16.

4. **Database**
   Migration of our offline database from MySQL to Oracle continues to be a long and painful saga and is still not yet complete. Dennis keeps a log of progress at http://home.fnal.gov/~dbox/minos/status_reports.html. He is currently trying to get everything working with Oracle 10. We were hoping to use the new IEEE floating point format in Oracle 10g but we have run into problems with truncation of data in the Perl DBI and so we are backing off to the internal Oracle representation – we use the Perl DBI to implement the whole check-summing mechanism that validates the contents of the database. Dennis has fixed the handling of exceptions so we can now hopefully debug some of the other issues that are still pending.
   We had a discussion about whether to drop the multi-layered system in favor of the MySQL and Oracle interfaces that are now available in ROOT. It turns out that this would require additional development on the ROOT end as the interface does not support binary data being returned from queries, it is all converted into ascii. This has significant performance implications. There is interest in the ROOT community on working on this so we will re-evaluate this option at a later date. We have strengthened many the SAM passwords in the Oracle DB to comply with the new security rules. In addition any new passwords must meet the new complex password rules.
   We will be upgrading the production server to Oracle 10g on March 16.

5. **DocDB**
   John Urish is making good progress on figuring out how to import the existing 900 NuMI notes into DocDB. The numbers have been reserved but the nature of the entries in the existing database and the complexity of the DocDB tables make this tricky. We have

complained about the outages caused by inadequate testing of changes/upgrades. I hope this will improve in the future.

6. **User Consulting**
   There is ongoing support of users which has the tendency to slow down progress on projects.

MINOS Offline Software

Most of the offline software is written by university collaborators. However we are responsible for certain of the infrastructure and simulation. This work is done primarily by Robert Hatcher. Much of the work in the last three months has been in support of the analysis of the first $1x10^{20}$ protons on target.

1. **Geometry**
   There have been some ongoing discussions with people in the ANL group on alignment and geometry. Some improvements were made to the existing geometry code.

2. **Detector Simulation**
   Work continues to progress, albeit more slowly that we would like, on the new C++ simulation interface. There was a joint meeting held between MINOS and NOvA regarding joint Monte Carlo tools.

3. **Box opening support**
   The box opening was delayed as the collaboration did not feel we were ready in Jaunary. Robert has continued to be involved in the simulation-related activities. This work continues to distract him from working on the new C++ simulation interface which continues to make slow progress. There have been a lot of meetings associated with this work.

4. **Restructuring of the offline setup**
   We have recently started creating regular snapshot releases of the code that are more stable than development but more recent than the frozen releases. We will move to nightly builds of development against the head of the ROOT CVS (the MINOS code is already built nightly but against a frozen version of ROOT). To facilitate the use of these new snapshot releases we have restructured the offline setup to use UPS.

5. **Support**
   A large amount of effort goes to user support. This has the usual impacts on on-going projects.

Resources used (budget, effort)

I summarize the percentage effort for the last three months. March is not yet available but I imagine it will look similar to February.

MINOS Computing

	Jan 06	Feb 06	Mar 06
E. Buckley-Geer	20	30	
Art Kreymer	45	85	
John Urish	50	90	

MINOS Offline Software

	Jan 06	Feb 06	Mar 06
Robert Hatcher	80	95	

MINOS Database Development

	Jan 06	Feb 06	Mar 06
Dennis Box	10	25	

Requested Budget for FY06
Amounts in K$.

Task	Task Code	Budget	Obligated	Comment
MINOS-DISK	50.1.06.04.05	30		Dcache disk and analysis disk
MINOS-CPU	50.01.06.04.06	112		Farm and analysis CPU
MINOS-COMPUTING	50.01.06.04.01	23		General operating
TAPES-MINOS-OP	50.01.10.13	22.5		Tapes

Schedule/Milestones

Implement file concatenation for production output. We want to implement this before the next major reprocessing pass which will probably start in a month or so.

Import existing Monte Carlo files into SAM. We are behind on this one. Get remote sites to produce necessary metadata for new files.

Finish migration of the existing NuMI notes into DocDB. The same tools to migrate the Neutrino Factory/Muon Collider notes into DocDB also.

Complete migration of offline database to Oracle. Your guess is as good as mine. Last milestone was June 2005 which we clearly missed by miles. We do not have a new milestone.

Try to make more progress on the C++ simulation. Robert needs to become a hermit for a few months but this is of course a bit difficult!